REPORT DOCUMENTATION PAGE

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
02/15/2006	Final	03/31/2001 - 09/30/2005

4. TITLE AND SUBTITLE

Improving 30-day Forecasts of Great Lakes Ice Cover: Phase III 30-day Spatial Forecasts

5a. CONTRACT NUMBER

N00173-01-P-6203

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Austin, Meg

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University Corporation for Atmospheric Research
PO Box 3000
Boulder, CO 80307

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Stennis Space Ctr, MS 39529-5004

10. SPONSOR/MONITOR'S ACRONYM(S)

NRL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

While many previous studies have highlighted important ice-atmosphere relationships in the Great Lakes, there remains a disconnect between the science and an objective 30-day forecast. In particular, most previous studies focused on temporally consistent (e.g. winter ice cover and winter mean temperatures) analyses, which are not applicable for forecasting. Additionally, most studies demonstrated only basin-scale linkages (i.e. the Great Lakes a single entity), but forecasters need information on scales specific to individual lakes or ports. Thus, the major challenges facing development of an objective statistical 30-day ice forecast are related to a lack of studies concerning how sub-basin scale ice cover variations are associated with atmospheric and ice conditions in preceding months.

15. SUBJECT TERMS

Great Lakes
Data analysis
IDL Program

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Meg Austin
U	U	U	UU	14	19b. TELEPHONE NUMBER (Include area code) 303-497-8630

Improving 30-day forecasts of Great Lakes Ice Cover: Phase III

30-day Spatial Forecasts[1]

Research done by the
UCAR Visiting Scientist Program at the National Ice Center

Summary of Research Report

Principal Investigator: Ms. Meg Austin

Report Period: 3/31/01 through 9/30/05

Recipient's Institution: University Corporation for Atmospheric Research
Visiting Scientist Programs
P.O. Box 3000
Boulder, CO 80307-3000
Voice: 303-497-8630
FAX: 303-497-8668
Email: austin@ucar.edu

Order Number: N00173-01-P-6203

LONG TERM GOALS

The long-term goal of the University Corporation for Atmospheric Research (UCAR) Visiting Scientist Program at the National Ice Center (NIC) is to recruit the highest quality visiting scientists in the ice research community for the broad purpose of strengthening the relationship between the operational and research communities in the atmospheric and oceanic sciences.

The University Corporation for Atmospheric Research supports the scientific community by creating, conducting, and coordinating projects that strengthen education and research in the atmospheric, oceanic and earth sciences. UCAR accomplishes this mission by building partnerships that are national or global in scope. UCAR's goal is to enable researchers and educators to take on issues and activities that require the combined and collaborative capabilities of a broadly engaged scientific community.

---

[1] With a brief re-cap of Phase I and II

OBJECTIVES

The objectives of the UCAR Visiting Scientist Program at the NIC are:
- Manage a visiting scientist program for the NIC Science Center in support of the mission of UCAR.
- Provide a pool of researchers who will share expertise with the NIC and the science community.
- Facilitate communications between the research and operational communities for the purpose of identifying work ready for validation and transition to an operational environment.
- Act as a focus for interagency cooperation.

The NIC mission is to provide worldwide operational sea ice analyses and forecasts for the armed forces of the U.S. and allied nations, the Departments of Commerce and Transportation, and other U. S. Government and international agencies, and the civil sector. The NIC produces these analyses and forecasts of Arctic, Antarctic, Great Lakes and Chesapeake Bay ice conditions to support customers with global, regional and tactical scale interests. The NIC regularly deploys Naval Ice Center NAVICECEN Ice Reconnaissance personnel to the Arctic and Antarctica in order to perform aerial ice observation and analysis in support of NIC customers. NIC ice data are a key part of the U.S. contribution to international global climate and ocean observing systems.

APPROACH

The UCAR Visiting Scientist Program works with participating Federal agencies to recruit scientific visitors and recent PhDs who are interested in conducting applications-oriented research and product evaluation of relevance to UCAR and the NIC ice-monitoring mission. The UCAR visiting scientists are a source of expertise for the NIC as well as mentors to the recent PhDs.

WORK COMPLETED

Great Lakes Forecast System Version 1.0
Research by Sheldon Drobot
Colorado Center for Astrodynamics Research
University of Colorado–Boulder
drobot@colorado.edu
http://ccar.colorado.edu/~drobot

Background

The benefits related to skillful 30-day forecasts of Great Lakes ice cover are clearly significant, given that over 1.4 billion metric tons (with an estimated value of $200 billion) have been transported along the Great Lakes since 1959. The U.S. Navy/NOAA National Ice Center (NIC) is responsible for issuing 30-day forecasts of ice conditions on the Great Lakes. Based on historical findings that variations in the ice cover are related to air temperatures, the NIC product currently utilizes estimates of freezing degree days (FDDs; a temperature proxy), analyst knowledge, some teleconnection data (NAO and ENSO), and analogues of previous years to issue a 30-day forecast. Recent findings indicate that variations in the ice cover are associated with other teleconnection patterns, such as the TNH and the PNA, and that these are valuable in forecasting ice cover at a 30-day interval. These recent findings offer hope that more skillful forecasts are possible, since most long-range forecasts rely on teleconnections in some way.

While many previous studies have highlighted important ice-atmosphere relationships in the Great Lakes, there remains a disconnect between the science and an objective 30-day forecast. In particular, most previous studies focused on temporally consistent (e.g. winter ice cover and winter mean temperatures) analyses, which are not applicable for forecasting. Additionally, most studies demonstrated only basin-scale linkages (i.e. the Great Lakes a single entity), but forecasters need information on scales specific to individual lakes or ports. Thus, the major challenges facing development of an objective statistical 30-day ice forecast are related to a lack of studies concerning how sub-basin scale ice cover variations are associated with atmospheric and ice conditions in preceding months.

Purpose of this Report

This report summarize phase III of the project, with a brief re-cap of Phases I and II. In phase I, statistical 30-day forecasts of lake-averaged ice cover were developed using preceding sea ice cover, NWS station data, and atmospheric teleconnection data (see Assel et al., 2004a, b for complete details). In brief, the following results emerged:

1. With empirical data available 30 days prior to a prediction, utilizing climatology[2] provides the lowest error for the 1 January Lake Ontario ice cover prediction, while the anomaly propagation model supplies the lowest error for the 1 March forecast for Lakes Superior, Michigan, and Huron, and the observed linear regression model provides the lowest error for the remaining 11 predictions.

2. For predictions of 1 January ice cover on 1 December, the mean November Tropical Northern Hemisphere teleconnection index provides the lowest prediction error, excepting Lake Ontario, where the climatological 1 January ice cover is a superior prediction. Including the November North Atlantic Oscillation index is helpful in reducing the error for Lakes Michigan, Huron, and Erie. For predictions of 1 February ice cover on 1 January, the 1 January mean lake-ice cover typically is the most important predictor, except for Lake Ontario, where the December East Atlantic-Western Russia teleconnection index is more valuable. For predictions of 1 March ice cover on 1 February, the 1 February ice cover generates the lowest error for all lakes

---

[2] The mathematical and conceptual foundations for the climatology, anomaly propagation, and regression models are discussed in section 3.

except Erie, where the accumulated freezing degree days over Lake Erie in January provides a lower error prediction.

3. If perfect forecasts of the upcoming month's freezing degree-days were available, the forecast equations for most months would need fewer parameters, and the error would be lower for all months. This suggests that as numerical weather models improve their accuracy of 30-day forecasts, analysts should consider utilizing these predictions more rigorously in the 30-day ice forecast.

While the results from Phase I showed promising statistical relationships, they are limited because they only predict one value per lake, while the typical Great Lakes forecast product is a spatial map (see Box 1 for a description of the 3 spatial scales used in this project). In phase II, statistical 30-day forecasts of freezing-degree-days were developed for National Weather Service Stations located on the banks of the Great Lakes. The following results emerged:

1. With empirical data available 30 days prior to a prediction, utilizing the regression models provides the lowest error for 17 of 18 stations for Dec FDD forecasts, 17 of 18 stations for Jan FDD forecasts, and 14 of 18 stations for February FDD forecasts. The climatological model is superior at 1 station for December and January FDD forecasts, while the historical regression model is superior at 4 stations for March FDD forecasts.
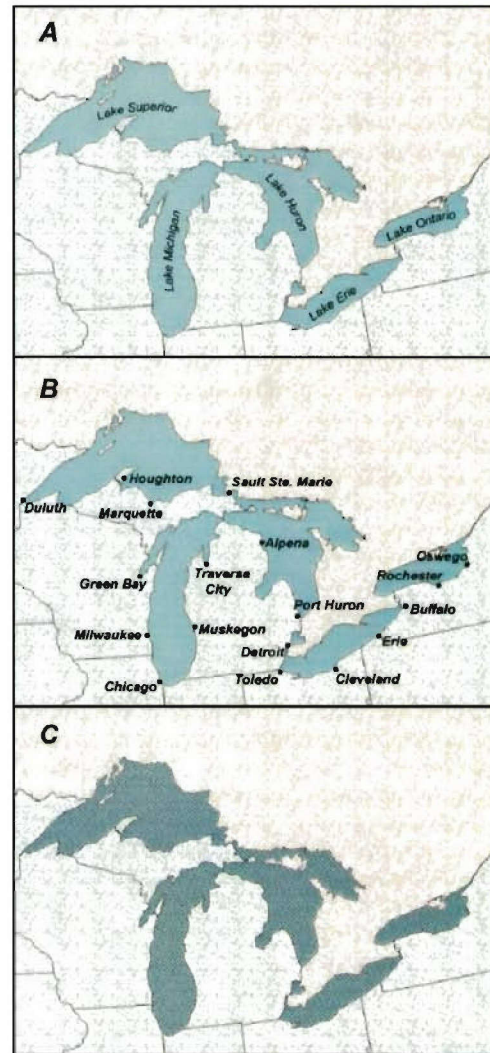

Figure 1. In phase I, forecasts were developed for each Great Lake (5 total); in phase II, forecasts were developed for each NWS station surrounding the Great Lakes (18 total); in phase III, forecasts are developed for cells (pixels) within each Great Lake.

2. Overall, the regression models provide relatively skillful forecast models in December, but there value is diminished in January and February. The results further emphasize the difficulty in predicting atmospheric conditions 30-days in advance.

Phase III Methods

Results from the preceding two phases clearly indicated the value of regression models for Great Lakes ice forecasting. Therefore, rather than simply repeating the methods from the first two phases, the approach is slightly altered in phase III, as shown in Figure 1.
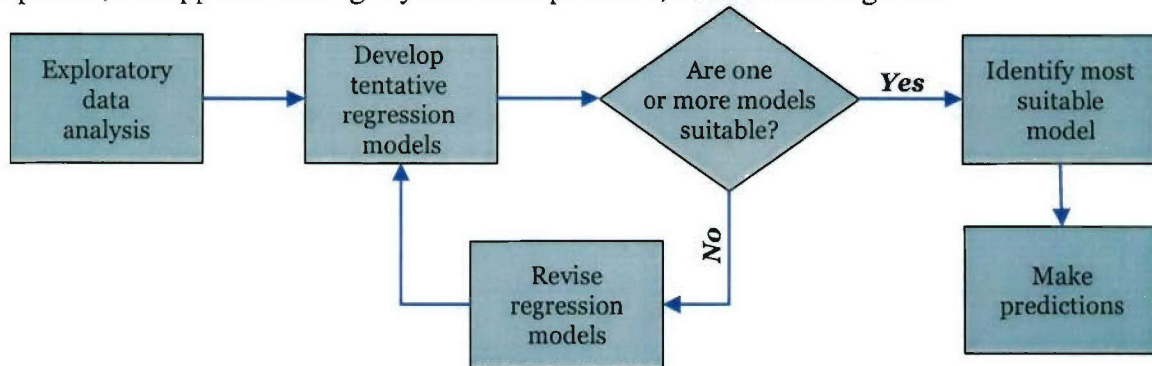

Figure 1. Schematic of data analysis framework for Phase III. See text for description of each phase.

In the **exploratory data analysis** phase, the 1973–2002 ice covers from Jan. 1, Feb.1, and Mar. 1 were correlated with Nov., Dec., and Jan. (respectively) teleconnection indices at each grid cell to highlight spatial regions where there are significant correlations. If a significant correlation was noted between the ice cover and the teleconnection index, then the teleconnection index was used in model development.

To **develop tentative regression models**, the all-subsets technique was used with the top 5 variables (i.e., teleconnection indices, FDDs, previous ice covers) that were most highly correlated with the ice cover for the month of interest. The all-subsets technique computes a linear regression for all combination of 1- through 5- variables, for a total of 120 possible regressions.

Determining whether **one or more models were suitable** involved testing whether the overall regression relationship was significant via the F-statistic. This is a common measure used for all regression models and it assesses whether the explained variance of the model is significantly different from 0. The level of significance was set to 0.05. Models that failed to explain any of the variance in the ice covers were deemed not suitable.

To **identify the most suitable model**, all suitable models were evaluated with a cross-validation scheme, whereby 80% of the data were used to develop the models and 20% were retained to test the MAE. This was repeated 100 times per model, so in effect, statistics were generated for 120 models and 100 iterations of each model. This cross-validation is necessary because independent data are needed to get a better estimate of model accuracy – if the models were evaluated with data used in developing the regression models, they would have an unrealistically low error.

Results
Exploratory Data Analysis

Correlation analyses between mean monthly November teleconnection indices and January 1 ice cover revealed several key teleconnections (Figure 2). The November East Atlantic/Western Russia (EAWR) index is significantly correlated with Jan 1 ice cover in the Eastern portion of Lake Superior, part of Lake Michigan, and much of Lake St. Clair.

In comparison, the East Pacific/North Pacific (EPNP) shows strong correlations along the shores of all Great Lakes, and is strongly correlated with ice cover over much of Lake Erie. The Pacific North American Anomaly (PNA) is even more strongly correlated with January 1 ice cover over much of Lake Huron, and there also are strong correlations in each of the other lakes.

In comparison, the Polar Eurasian (POL) index is significantly correlated in Lakes Huron and Erie primarily, and the Scandinavian (SCA) index shows some correlation in southern Lake Michigan. Combined, these indices provide significant correlations over most of the Great Lakes; information which will be utilized for 30-day forecasts. Correlation analyses between mean monthly December teleconnection indices and
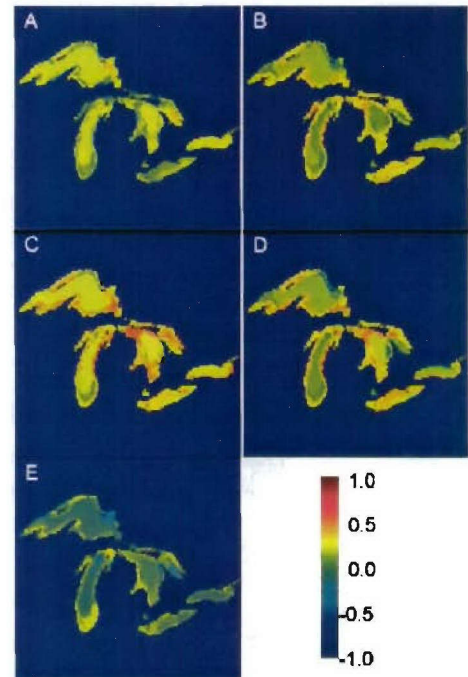

Figure 2. Correlation between January 1 and November teleconnetions for (a) EAWR, (b) EPNP, (c) PNA, (d) POL, and (e) SCA indices.
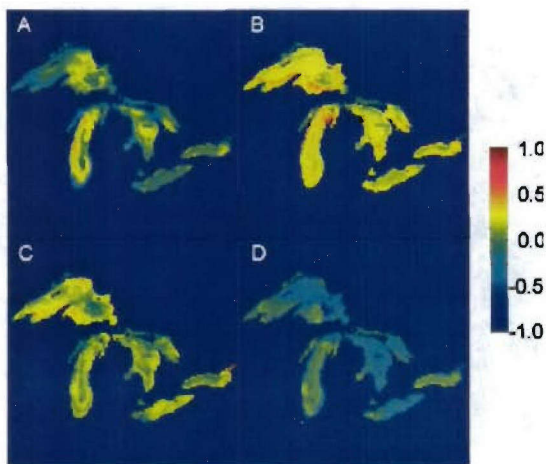

Figure 3. Correlation between February 1 ice cover and December teleconnetions for (a) PNA, (b)SCA, (c)TNH, and (d) WP indices.
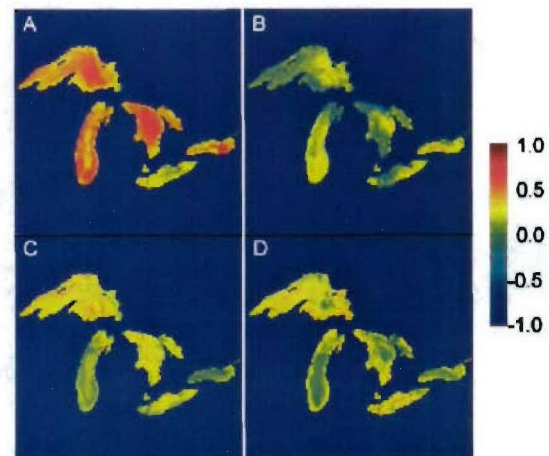

Figure 4. Correlation between March 1 and January teleconnetions for (a) EPNP, (b) PNA, (c) SCA, (d) TNH indices.

February 1 ice cover indicate results similar to January 1 (Figure 3). The SCA index the most strongly related correlation, although there are significant correlations in various portions of the Great Lakes for other indices as well. Correlation analyses between mean monthly January teleconnection indices and February 1 ice cover also demonstrate comparable significance with preceding months (Figure 4). The EPNP is by far the most significant variable, with strong correlations over every lake except Lake Erie. In comparison, the PNA shows significant correlation in parts of all five lakes and the SCA index is significantly correlated mainly in Lake Superior. The TNH index shows the strongest correlation in Lake Erie.

Model Results

The best model to forecast the January 1 ice cover included the following November variables:

- ❖ EAWR
- ❖ EPNP
- ❖ PNA
- ❖ POL
- ❖ SCA

Variations in these five parameters resulted in low MAE in the interiors of the Great Lakes (Figure 5a), but MAEs near 50% along the coasts. In comparison with climatology (Figure 5b), this model provided improvements mainly in Lake Erie, Lake St. Clair, and part of Lake Superior (Figure 5c).

The best model to forecast the February 1 ice cover included the following December variables:

- ❖ PNA
- ❖ SCA
- ❖ TNH
- ❖ WP
- ❖ and Jan 1 ice cover

Similar to January forecasts, variations in these five parameters resulted in low MAE in the interiors of the Great Lakes (Figure 6a), but MAEs near 30% along the coasts. Nonetheless, in comparison with climatology (Figure 6b), this model provided significant improvements in Lake Erie (Figure 6c).

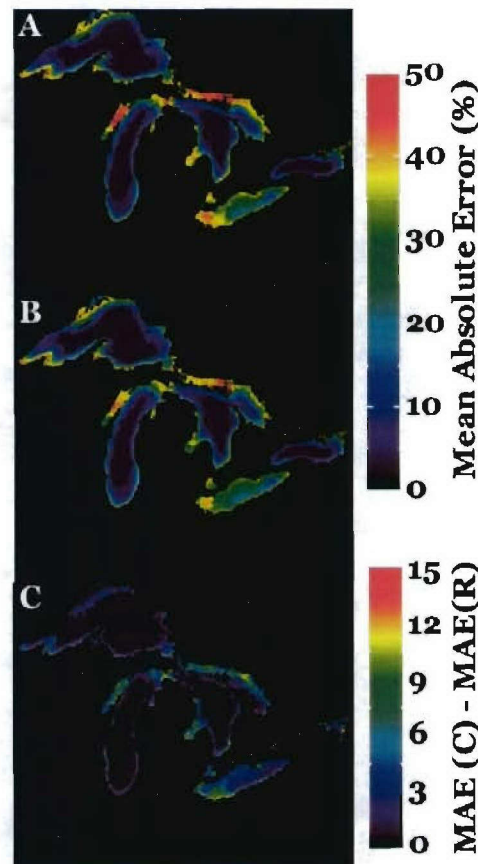The best model to forecast the March 1 ice cover included the following January variables:

Figure 5. Statistics for January 1 forecasts. Mean absolute error for (a) climatology forecasts, (b) regression forecast, and (c) difference in MAE between climatology and regression forecast.

- ❖ EPNP
- ❖ PNA
- ❖ SCA
- ❖ TNH
- ❖ and Feb 1 ice cover

Similar to January forecasts, variations in these five parameters resulted in low MAE in the interiors of the Great Lakes (Figure 7a), but MAEs near 30% along the coasts. Nonetheless, in comparison with climatology (Figure 7b), this model provided significant improvements in Lake Superior (Figure 7c).
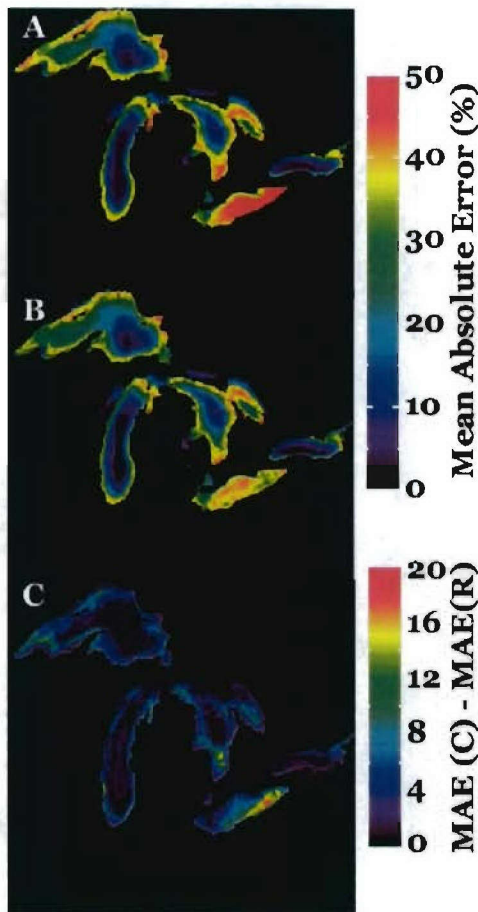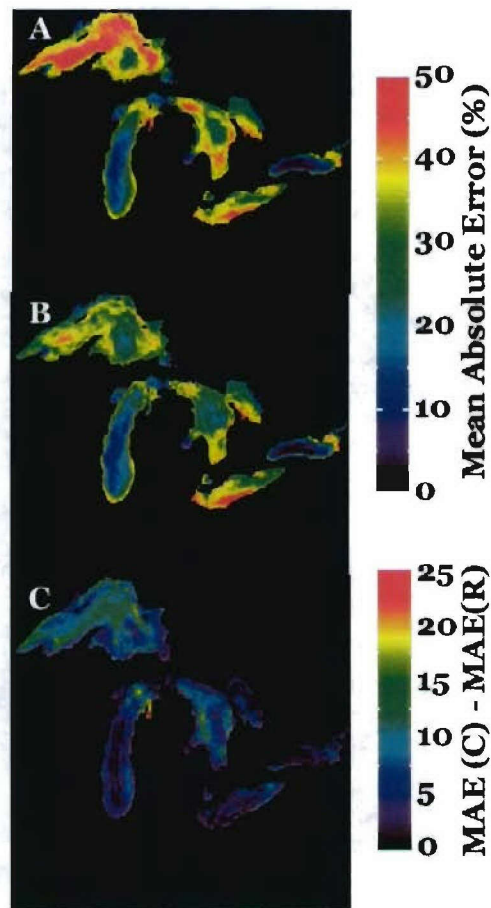

Figure 6. Statistics for February 1 forecasts. Mean absolute error for (a) climatology forecasts, (b) regression forecast, and (c) difference in MAE between climatology and regression forecast.

Figure 7. Statistics for March 1 forecasts. Mean absolute error for (a) climatology forecasts, (b) regression forecast, and (c) difference in MAE between climatology and regression forecast.

Setting-up and running the IDL program

This section discusses the steps needed to run the IDL program for spatial Great Lakes forecasts.

Step 1: Set up the files
- Anonymous FTP to ccar.colorado.edu
- Switch to pub/drobot/nic
- Copy all the files to F:\Great Lakes. You should have the main IDL program, Great_Lakes_Forecast_System.pro and a backup of this program, Great_Lakes_Forecast_System.bak. There also will be 8 .flt files for each month. The .flt means the data are floating point arrays in IDL binary format. You should never have to do anything with the files.
- Note: I have also placed the GLERL files used in creating the equations at pub/drobot/glerl_jan, pub/drobot/glerl_feb, and pub/drobot/glerl_mar. These are optional for downloading. You don't need them for future forecasts, but to test out the program and see how things work, you will probably want them.

Step 2: Load the program in IDL
- Either double-click on the Great_Lakes_Forecast_System.pro file in the Windows directory, or you can load IDL and then go File>Open and find Great_Lakes_Forecast_System.pro

Step 3: Compile the program
- Click on the compile button from the top menu of IDL (see figure below). You'll need to do this twice, because the first program (Great_Lakes_Forecast_System) calls other programs listed below the Great_Lakes_Forecast_System (e.g., texBox, colorbar) that need to compile before Great_Lakes_Forecast_System will work. If I had put those other programs first, then you would not need to do that, but I wanted to put the main program (Great_Lakes_Forecast_System) first in case you want to see what the code is doing. So that's a little odd, but just hit that button twice and you'll be fine.
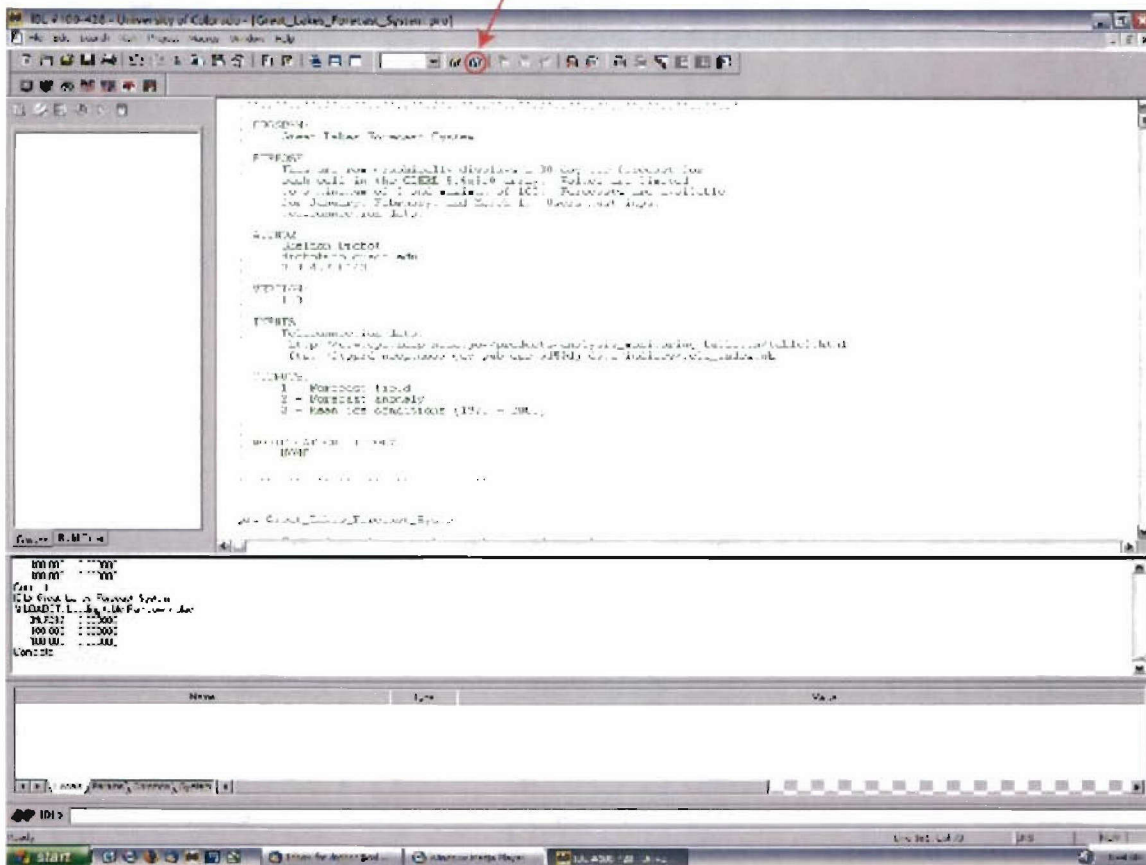
First click here to compile the program
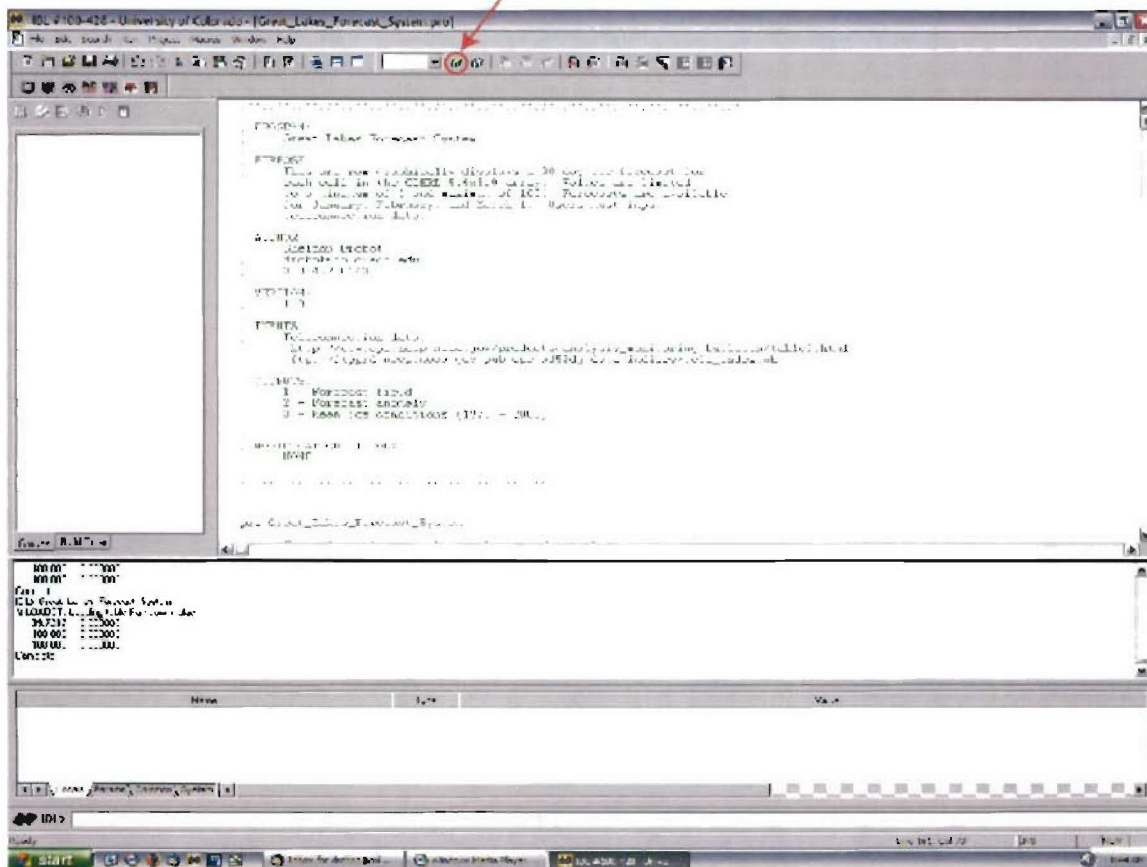
Step 4: Run the program
- Now click on the button to the right of the previous one and the program will run.
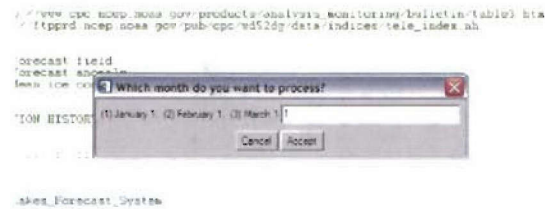
Further Notes on the program:

- After you run the program, the first thing that happens is a pop-up welcome (note: I am trimming the images to save space). This just gives you the current version and my contact info. Click 'OK' to continue.

- After clicking 'OK', the program will ask you which month you want to forecast. Valid answers are 1 for January, 2 for February, and 3 for March. You need to click inside the white box to activate the cursor. Hit 'Accept' when you are ready to move on, or cancel if you want to quit. If you enter 0 or something greater than 4, you'll get a pop-up error message and the program will quit.

/ /www cpc ncep.noaa gov/products/analysis_monitoring/bulletin/table3.html
/ ftpprd ncep.noaa gov/pub/cpc/wd52dg/data/indices/tele_index.nh

orecast field
forecast anomaly
mean ice co...

- After selecting a valid month, a pop-up will confirm which month you selected and tell you that you next need to enter 5 (for January) teleconnection values and where to get the data. If you select February or March, it will ask for 4 teleconnection values and the preceding month's ice cover. Click 'OK'
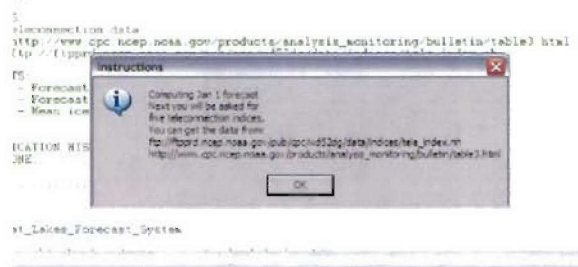
- The teleconnection values are entered the same way as the month was previously; you just click into the white box and enter the value

- If you are processing February or March, then you'll get a screen asking where you can find the previous month's ice data. This is the GLERL ice data from the first day of the last month. It can be named anything you want. This is probably the step where an error might occur. Right now, the program is looking for the GLERL-formatted data, so it is important to make sure the data are formatted correctly, which is '(1x,516i2)'. If you run into trouble here, we can try and load it in free-format, but the way it is set-up right now is better.
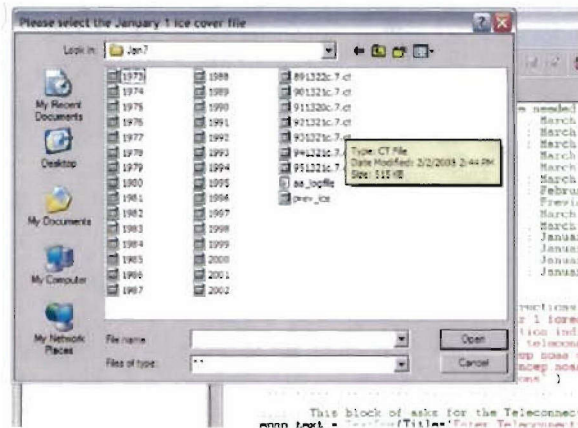
- Next the program will run the regression equations and then give you two maps. The top right is the mean absolute error. This tells you where you can be most confident in the results, with smaller numbers being better. The second map, in the lower right, is the actual forecast. The program asks you to save this file (which you can cancel out of and keep going if you want). The file will be saved as a .PNG, based on whatever name you want to give it.
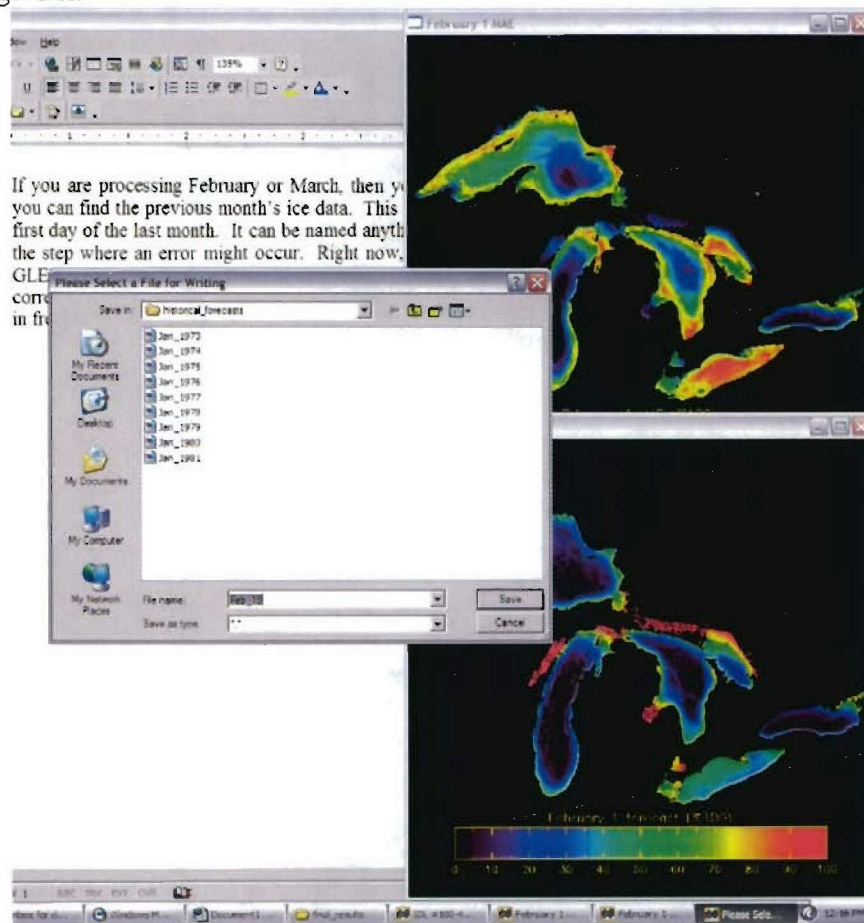
- After you respond to either saving or not saving, the final two maps are drawn. In the top left is the 1973-2002 climatological mean values, and the lower left is the forecast-climatological mean values. This lower left image does not render terribly well in some cases, so that is something I will fix for version 1.1!